DOCUMENT RESUME

ED 084 223 SP 006 971

AUTHOR	Buck, James J., Comp.; Parsley, James F., Jr., Comp.
TITLE	The Way We See It: A Survey of Teacher Evaluation Policies and Practices Operant in the State of Washington.
INSTITUTION	School Information and Research Service, Seattle, Wash.
PUB DATE	Mar 73
NOTE	19p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Evaluation Techniques; *Instructional Improvement; Lesson Observation Criteria; Performance Criteria; School Districts; *School Surveys; *Teacher Evaluation; Teacher Rating
IDENTIFIERS	Washington

ABSTRACT

This paper assesses teacher evaluation policies and practices operant in First Class School Districts in the state of Washington. The data, submitted from mailed questionnaires, deal with a district's assessment policy and specifically include concerns relating to evaluative personnel, procedures, and an evaluative instrument or model. Approximately 60 percent of the school districts responded. The following conclusions were made: a) most districts utilized a district-wide evaluation model; b) the purpose of teacher assessment was instructional improvement; c) principals were the primary evaluators, with peer evaluators increasing; d) observation was the most frequent method for evaluation, with the uses of performance objectives and self-evaluation techniques increasing; e) personal characteristics and instructional skills were the criteria used in evaluation; f) the rating instrument and the conference were the most common assessment forms; and g) most districts are now reviewing or revising evaluation programs. Recommendations for improved programs are made. A two-page bibliography is included. (BRB)

SIRS SCHOOL INFORMATION AND RESEARCH SERVICE

APR 10 1973

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 084223

BULLETIN

March 30, 1973 1972-73: 3

IN THIS ISSUE:

THE WAY WE SEE IT:

A Survey of Teacher Evaluation Policies and Practices Operant in the State of Washington

Compiled by:
James J. Buck, 1971-72 Cooperative Center Fellow
and
James F. Parsley, Jr., Director
Cooperative Washington Education Centers

FILMED FROM BEST AVAILABLE COPY

SP 006 971

200 EAST UNION AVE.—OLYMPIA, WASH. 98501 PHONE: 943-5717

This study, sponsored by the Cooperative Washington Education Center, attempts to clarify the highly complex problem of the assessment and evaluation of teachers. Although it is comprised mostly of non-empirical data and lacks statistical sophistication, it can be a valuable index of common practices in the daily operation of school districts.

The study further examines how the intent of Washington law RCW 28A. 72.030 is being carried out by school districts with regard to teacher assessment and evaluation.

Finally, the study cites some of the implications of teacher assessment and evaluation, and offers some practical suggestions in developing your own evaluation programs.

The Editor

THE WAY WE SEE IT: A Survey of Teacher Evaluation Policies and Practices Operant in the State of Washington

CHAPTER I

INTRODUCTION

The past decade has seen education enveloped in myriad crises causing dissipation of public confidence in the schools. Disparate views concerning the role of public education ranged from those who were accustomed to viewing the educational institution as a societal godsend with the antithetical view increasingly being expressed by doubting the schools' ability to meet pending social needs. The current financial crisis confronting public school educators has become a prominent national concern and has in part led to a mandate to devise assessment techniques which yield the type of data which meaningfully enhances public knowledge concerning the effectiveness of current programs and practices. Thusly, this concern has forced educators to consider the assessment problem and subsequent evaluative procedures which have never before been widely used by the education community.

Apart from the obvious and immediate application to instructional program effectiveness the mandate to become increasingly accountable has extended to teacher effectiveness. Nationally the aforementioned trend seems incontrovertable while in our own State of Washington recent state law has sought to clarify the responsibility of assessing teacher effectiveness:

> Every board of directors, in accordance with procedure provided in RCW 28A. 72. 030, shall establish an evaluative criteria and procedures for all certificated employees. Such procedure shall require not less than annual evaluation of all employees. New employees shall be evaluated within the first ninety calendar days of their employment. Every employee whose work is judged unsatisfactory shall be notified in writing of stated areas of deficiencies along with recommendations for improvement by February 1st of each year. A probationary period shall be established from February 1st to April 15th for the employee to demonstrate improvement.

While public insistence has intensified the need for accountability procedures, efforts at formulating evaluative tools had already begun as a direct result of change in education. The proliferation of teaching strategies, the increased monetary appropriations through federally administered Title programs, and the wealth of educational materials decidedly makes education in the State of Washington a non static entity. In this transitional environment, decisions must be evaluated, organizational procedures must be appraised, and materials must be tested. Change, if it denotes progress, must be accompanied by evaluation, out of which should emerge the basis for all subsequent change.

I. THE PROBLEM

Statement of the Problem

The present study attempts by means of questionnaire to survey teacher evaluation policies and practices operant in First Class Districts in Washington State, and to report various concerns arising from these evaluation programs as currently conducted.

Purpose of the Study

In October, 1971, district coordinators for the Cooperative Washington Education Centers assembled at Central Washington State College to discuss common educational concerns including that of teacher accountability. As a result of this meeting, the Cooperative Center commissioned this survey in the hopes of assisting school districts in their development of more effective teacher evaluation programs. The purpose of the survey is to enable school districts in Washington State to learn of diverse evaluative methods currently employed, the primary intent being to provide a background on which districts might revise their evaluation programs as they deem advisable. It is anticipated further, that ideas gleaned from the survey will ultimately effect instructional improvement by encouraging school districts to devise evaluation programs conducive to professional growth on the part of the teacher.

II. DEFINITION OF TERMS

For the purpose of this study the following terms have been defined.

First Class School District

A First Class School District is a school district having in excess of 10,000 population as shown by regular census or any other evidence proved acceptable by the local Intermediate School District.

Cooperative Washington Education Center

The Cooperative Washington Education Center is an organization comprised of twenty-four institutional participants who benefit from a spirit of mutual cooperation and concern. The Cooperative Center in its attempts to promote needed change and to reduce the duplication of efforts in school districts, has developed means of providing for an exchange of innovative ideas. The Educational Practices Inventory, a compilation of innovative programs operant in the member centers, and the annual in-service workshop allow for a sharing and dissemination of applicable practices. Through a shared services program, districts may request, at a no-cost basis, other district personnel possessing expertise in curricular content areas and/or in educational administration. The Principal's Leadership Training Project, initiated in 1970, assists participating principals in developing administrative competencies essential to their positions. Representing indirectly one-quarter of the State's pupil enrollment, the Cooperative Washington Education Center has from the onset been concerned with the dissemination of innovative programs and practices through hosting an annual statewide education conference and publishing a number of monographs, pamphlets, and newsletters.

Evaluation Programs

Numerous districts reported their evaluation programs of teachers under current review and/or revision. The data in this survey reflects policy initiated prior to 1972 with one exception: if a district indicated an experimental or pilot program had been implemented and its formal adoption by the district was probable, then this new program was included in the study rather than the district's previously stated policy.

Teacher Accountability

Teacher accountability refers to the assessment of teacher performance as a causal factor in the pupils' achievement of desired objectives in light of resources provided.

III. LIMITATIONS OF THE STUDY

This study was dependent upon the information and data submitted by the First Class Districts surveyed. Conclusions are drawn from the approximate sixty per cent of all First Class Districts responding to the survey. This study describes these districts' policies governing teacher evaluation. It does not state how assiduously these policies are enforced or carried out, nor does it report on the relative success of the various evaluation programs.

CHAPTER II

REVIEW OF THE LITERATURE

Few topics have been as thoroughly researched as the subject of evaluating teacher effectiveness. As Biddle noted in his preface to Contemporary Research on Teacher Effectiveness:

> Probably no aspect of education has been discussed with greater frequency, with as much deep concern, or by more educators and citizens than has that of teacher effectiveness--how to define it, how to identify it, how to measure it, how to evaluate it, and how to detect and remove obstacles to its achievement (5:4).

Yet the research on teacher effectiveness has produced deplorably few results. The persistent problem of determining what teaching leads to what learning continues to plague administrators and teachers alike. Years of research and experience have left many educators with serious doubts concerning the efficacy of teacher evaluation efforts. Only by reviewing the literature can one appreciate the burden that has fallen to education.

Prior to and during the 1930's few educators denied that one could competently judge a teacher's proficiency by observing his action in the classroom. Consequently, rating scales were devised to assist supervisory personnel in measuring teacher efficiency. These rating scales typically consisted of items relating to the teacher's social relations, instructional skills, personal characteristics, and professional qualifications. Researchers, meanwhile, were consistently finding that the criteria utilized in the rating scales were exceedingly arbitrary and vague. Reavis and Cooper (1945), after reviewing rating devices of 103 school systems, concluded: "Ratings appear to be invalid, then as a comprehensive measure of either general or specific teacher ability" (5:61). Anderson (1954) confirmed the earlier conclusions drawn by Hellfritzch (22) and Lins (26) by testing eight different evaluative criteria: ratings by the principal, by peers, by pupils, by the teacher herself, by an outside agency, and by pupil achievement scores. He declared in his summary:

> No correlation appreciably different from zero was discovered between the evaluations of the teachers on the different rating scales and the evaluations based on the achievements of their pupils in the subject matter areas (2:68).

Numerous other studies substantiated Anderson's conclusion. Statements similar to Rabinowitz and Travers, "No teacher is more effective than another except as someone so decides and designates. . ." (17:37) abound in the literature.

Walker, almost twenty years before Anderson, described the core difficulty in evaluating teacher effectiveness, and, in so doing, exemplified the prevalent attitude of that period:

> The lack of an adequate, concrete, objective, universal criterion for teaching ability is thus the primary source of trouble for all who would measure teaching. One typical method of attack used in rating scales is to compile a list of broad general traits supposedly desireable for teachers, with respect to which the rater passes judgment on each teacher. . . Even when the scale is made quite specific, relating not to general traits but to concrete procedure, the fundamental difficulty remains, that there is no external and generally accepted criterion against which the scale can be validated to establish the significance of its items (16:602).

This passage implies that there is some variable for measuring teacher effectiveness which applies to all teachers regardless of level, situation, or subject. This single criterion has yet to be defined, allowing Morsh and Wilder (1954) to confidently report:

> No single, specific, observable teacher act has yet been found whose frequency or percent of occurrence is invariably and significantly correlated with student achievement (13:1423).

Gage, in his comprehensive analysis of this 'criterion of effectiveness' paradigm, explained that, even if criteria were found to be related to teacher effectiveness, there would be disagreement as to the ultimacy of the criteria. For example, which of these criteria would contain more weight: the teacher's effect on pupil achievement of current educational objectives; or the teacher's effect on pupil achievement in subsequent schooling? As decades passed, the search for a universal criterion gradually ebbed. Gage, in his Handbook of Research on Teaching, eulogized such efforts perhaps most fittingly:

> Research by this paradigm (criterion of effectiveness) has been abundant; hundreds of studies . . . have been made. In the large, these studies have yielded disappointing results: correlations that are nonsignificant, inconsistent from one study to the next, and usually lacking in psychological and educational meaning (15:18).

This vein in (what we might term) the "Accountability Mountain" had been mined out and abandoned, with researchers looking elsewhere for more profitable diggings.

Objectivity continued to be the prospector's pick, and the introduction of merit-pay proposals served to intensify the urgency for developing objective standards to measure teacher performance. Once again, exhaustive state-wide studies achieved only minimal success in defining behaviors which were reliably associated with pupil achievement. In his state-by-state analysis, McPhail stated:

> As indicated in the literature, several states (Delaware, Tennessee, Mississippi, New York, North Carolina, and Utah) have been concerned and involved in attempts to evaluate teachers by factors in addition to degree and experience. . . most have proven unsuccessful in determining facts of and/or procedures that have proven to be workable; therefore, in most instances, they have been discontinued (27:356).

Another approach used to measure teacher effectiveness has been that of comparing some teacher characteristic or personality trait with an outcome variable. Generally, researchers found few correlations between teacher variables (sex, age,

socio-economic background, academic achievement, marital status, intelligence, and voice quality) and teacher effectiveness. Ryans conducted an indepth study of the characteristics of teachers, but failed to arrive at any variable which was correlated significantly with effectiveness in the classroom. He attributed the reason to the wide variations in defining "good teacher." As he explained:

> A person's concept of a good teacher seems to depend on (a) his acculturation, his past experience, and the value attitudes he has come to accept, (b) the aspects of teaching which may be foremost in his consideration at a given time, and (c) characteristics of the pupils taught (38:370-371).

Still another approach hypothesized that teacher attitudes and values were related to teacher effectiveness, and could serve also as adjunct predictors of teacher performance. Attitudinal inventories and personality tests, especially the Minnesota Teacher Attitude Inventory (MTAI) and the Minnesota Multiphasic Personality Inventory (MMPI), attained notable prominence, although numerous critics (Munro (33), and Hedlund (17)) attacked their validity in predicting successful teachers. Both Rabinowitz and Coleman found the MTAI susceptible to "faking" and found that the "subject with some knowledge of the viewpoint endorsed by the selection agency could in most cases reflect this viewpoint in responding to the MTAI" (17:520).

Heil and Washburne (21), however, asserted there were identifiable types of teachers, as classified by the Manifold Interest Schedule, and that the different types (three in their study) have different effects on the children they teach. For example, they discovered that pupils made the most progress in math and science under the "Type A" (turbulent, impulsive, and variable) teacher than under the other two types. Similarly, Broudy reported that teachers who scored high on the MTAI tended to be more effective with students who are less cognitively oriented, and suggested that the teacher be paired with the pupils with whom she might be most effective (9).

With the extreme paucity of solid evidence supporting evaluation of teacher effectiveness, the American Association of School Administrators, the Department of Classroom Teachers of NEA, and the National School Boards Association acknowledged in a joint publication.

> The notion of the "good teacher" so basic to study of teacher effectiveness turns out to be almost as vague and diffuse as the range of human experience relative to teaching (17:4).

Increasingly, researchers concentrated on observing the interaction between teacher and pupil in the classroom. As Mitzel (1960) cautioned,

> In considering both teacher behavior and student behavior as process criteria it became clear that neither of them should be studied in isolation from the other. The interaction between them appears to be the dominant aspect of the whole process of learning (31:1484).

Hosts of interaction-behavior indices appeared as researchers experienced a renewed enthusiasm and optimism which permeated the decade of the sixties. Flanders' Interaction Analysis Procedure (14), Mork's Verbal Reaction Behavior Log (32), and Ober's Reciprocal Category System (34) are examples of instruments designed to record teacher and pupil behaviors, especially verbal behaviors, and to test the relationship between this behavior pattern and the attainment of educational objectives.

Anderson and Hunka, reviewing the literature on teacher evaluation, conceded:

> Studies by . . . Anderson, Brewer and Reed, Thelen, and Flanders have shown a relationship between certain sorts of teacher behavior and extent of growth by students along certain desireable intellectual and attitudinal criteria (1:75).

Despite the internal validity obtained in many interaction studies, researchers began to question how applicable and reliable the correlations would be if the same teachers were observed in a different situation. As Barr commented:

> Teaching does not take place in a vacuum; it takes place in a tangible situation . . . time has seen the emphasis shift from the teacher per se to the teacher in relation to the more important aspects of a situation (3:141).

Wick and Beggs concur by saying, "The question, 'What are the dimensions of an effective teacher?' cannot be answered without asking, 'In what situation?'" (43:194). Even after Turner and Fattu (1:76) concluded that the most economic and systematic way of conceptualizing the potential inexhaustible concept of teacher was professional problem solving, Flanders responded:

> The weight of the evidence from the studies reviewed suggests that problem solving performance is a measurable characteristic of teachers and that this characteristic changes under teacher preparation and experience but that its importance to teacher success <u>is largely contingent on the type of setting within which the teacher does his work</u> (author's emphasis) (13:1432).

Soar, critical of the restrictive view taken by past researchers similarly concluded:

> Once a multidimensional view of the teaching-learning process is accepted, it seems likely that a teacher may do some things that are not effective; or even that a given act may be effective in working toward one goal but not another; or that an act may be effective in teaching one child and ineffective with another (41:289).

Such doubts have caused educational researchers like Stufflebeam (42) and Guba (18) to propose evaluation strategy designed to foster decision-making on the part of the teacher. This approach implies that the teacher is an active member in the accountability process and is more apt to change if he participates in the decision to change. In short, such procedures call for the teacher to assess a particular situation and then determine what he can and should do to increase his effectiveness in that situation. Since supervisors assist in an advisory capacity, much of the negative, threatening aura surrounding many evaluative procedures and instruments is dispelled.

As a recent development, Gage suggests the use of 'microteaching' in determining the effectiveness of teacher performance. He explains the procedure as one of breaking down the global, complex variables explored previously, and selecting for study a single <u>facet</u> of teaching, such as explaining ability. The result, hopefully, would be the emergence of a set of lawful relationships between variables, something unattainable by those who attempted to study the whole, unanalyzable process of teaching (16). According to 'micro-teaching,' an instructor teaches a group of five individuals focusing on one aspect or facet of the role. Opponents to this approach have retorted sharply that a teacher does not teach five students in the same fashion as she does thirty. Also, the obvious artificiality of the situation, they say, precludes its adaptability for measuring teacher effectiveness.

Current accountability efforts involving performance contracting and voucher systems (notably Texarkana and Cherry Creek programs) have sparked additional research into evaluation. A proponent of accountability, Leon Lessinger, in Every Kid a Winner: Accountability in Education, elucidates the concept of 'educational engineering.' In this volume, he presents a detailed description of the Texarkana performance contract project, and elaborates on the procedures required in evaluating a school's effectiveness (25).

Contrastingly, Stephen Hencley (23), in a soon to be published work, presents a cogent examination of the impediments to accountability. He examines the philosophical, technological-economic, and political-legal deterrents to accountability. Hencley sees the intense concern for accountability as philosophically in conflict with the more open, affectively oriented, child-centered schools advocated by Silberman, Glasser and many others. In citing contemporary research on testing Hencley iterates a concern over the validity of product accountability as now measured. Hencley's and Lessinger's articulate treatments of the problem of accountability illustrate the wide disparity of views held by those in the educational community.

Conclusions

Literature relevant to the present study can be divided into three major areas, referred to by Mitzel as presage variables, process variables, and product variables. Each of these categories has been investigated testing its relation to teacher effectiveness. From the review it was concluded that:

1. Research linking presage variables with teacher effectiveness (referring to teacher characteristics, values, and attitudes) has produced few tangible results. Only very general trait designators such as "interested" and "sympathetic" have been found to correlate significantly with teacher effectiveness. Contradictory conclusions and findings abound in the literature. Disagreement over the definition of "good teaching" has contributed to the confusion surrounding the research, and has impeded efforts at identifying teacher traits related to effectiveness.

2. Though a recent development, the study of process variables has produced a greater research objectivity due to its emphasis on measuring pupil and teacher behavior. However, research results have been negligible, frequently positing only obvious patterns, for example, the positive relationship between teacher responsiveness to pupil ideas and subsequent pupil achievement. While researchers claim success in studying the relation of process to product variables, its time-consuming procedures are generally considered too impractical for implementation in the schools.

3. Product variables, referring to the effects on pupil behavior or achievement has received noteworthy attention by proponents of accountability. Research conducted on the effectiveness of performance contracting is as yet incomplete. Meanwhile, critics of accountability procedures have intensified their efforts in finding inherent weaknesses in test standards used to validate product achievement. The current trend in research has shifted from correlating presage and process variables with product variables focusing now on the relation of product variables themselves and teacher effectiveness.

CHAPTER III

METHODOLOGY AND RESULTS

The purpose of this chapter is twofold: first, to describe the sample selected and to clarify the procedures used in obtaining the data; secondly, to provide a description of the results derived from the information collected in this survey.

Methodology

This study intended to gather information and data pertinent to teacher evaluation programs conducted in first class school districts in Washington State.

A questionnaire, explaining the intent of the survey and specifying the data desired, was initially mailed to all first class school district superintendents in November and subsequently followed by a second mailing in April. Of the sixty-eight first class districts, forty districts (approximately 60 per cent) responded to the questionnaire. The data collected were handtabulated and the conclusions expressed in this study reflect the interpretations and views of the authors.

Results

Concern one: To what extent have school districts addressed themselves to the state mandate by adopting policy guidelines pertinent to teacher assessment?

In this survey, 95 per cent of the districts reported having adopted district policy regarding teacher evaluation. The remaining 5 per cent, although administering a teacher evaluation program, stated they had no formal district policy governing it.

Concern two: To what extent have individual schools within first class districts formulated their own teacher evaluation procedures rather than employ a district-wide model?

Data revealed that 82 per cent of the sample employed a district-wide evaluation model involving traditional methods of evaluation such as the rating checklist and specific comment instruments. Contrasted to the above finding, 10 per cent encouraged individual schools to develop their own evaluative programs. It was found that these schools, when developing teacher evaluation programs, tended to devise and implement procedures that varied from the typical principal rating used in most districts. These schools frequently instituted such evaluative procedures as video-taping and pupil assessment of the teacher. Also, data indicated a much higher incidence of peer evaluation as an alternative to principal evaluation in schools having developed their own assessment models.

An inverse relationship was found between a district's encouragement of diverse evaluative methods and its pupil enrollment or size. Thus, as a district increased in size, the likelihood of individual schools within that district developing their own evaluative procedures decreased.

Concern three: Since accountability emphasizes the attainment of educational objectives, this question was posed: to what extent have first class districts developed procedures for assessing achievement of performance objectives? Relatedly, what proportion of districts utilize these procedures in their teacher evaluation programs?

The development of procedures for assessing attainment of educational objectives was reported by 45 per cent of the districts. However, only 20 per cent had applied these procedures to their formal teacher evaluation programs. Another 25 per cent of the sample commented that their districts were presently engaged in the development of such procedures, but admitted the formulation of educational objectives was an arduous and complicated task. The remaining districts either responded negatively or omitted answering the question.

Concern four: Evaluation performs specific functions implying it is goal directed or purposive. In soliciting information, this survey inquired: what is the expressed purpose(s) of your district's teacher evaluation program?

Overwhelmingly, districts agreed that the primary purpose for their teacher evaluation program was instructional improvement. Compared to 88 per cent who selected this purpose, only 7 per cent cited legal compliance as the primary intent for their policy.

Several districts reported more than a single or primary purpose for their evaluation program. Secondary reasons given concerned legal compliance, personnel retention (tenure, promotion, dismissal), professional growth, and staff communication.

Several respondents, noting that the primary or "intended" purpose of instructional improvement was not necessarily the one accomplished, indicated that a possible incongruity existed between the evaluative procedures or practices and the designated purpose. Conversely, another administrator whose district policy, he believed, related most to legal compliance, remarked that evaluators were nevertheless encouraged to accentuate a positive attitude when in conference with individual teachers. A small minority of the sample, attempting to make the purpose of instructional improvement more clearly apparent, provided teachers with learning agents or co-helpers to assist in the development and achievement of performance objectives.

Concern five: Though teacher evaluation often assumes diverse forms and procedures, the evaluator is a necessary component. The following results were found when districts responded to this question: what personnel are primarily responsible in your district for assessing a teacher's performance?

The person most frequently reported to be assuming the role of teacher evaluator was the building principal. His involvement either as sole evaluator or as an evaluation team member was reported by 90 per cent of the districts. Only 5 per cent of the districts allowed teacher self-evaluation coupled with principal conference to suffice for the annual evaluation. The use of peer evaluators, completely excluding principal participation in the formal evaluation, was permitted in 5 per cent of the sample. Other districts encouraged peer evaluation as a supplementary practice for the teacher's benefit, but it did not replace or substitute for the district's standard procedure. Another means of teacher assessment, that of pupil evaluation, was conducted in only three districts, and in each instance only on a temporary, experimental basis, or in a specific school at the teacher's request.

Concern six: Evaluators follow specific procedural guidelines when assessing the effectiveness or the performance of a teacher. This survey addressed itself to this question. what procedures are utilized by evaluators to gather information that will serve as the basis for teacher assessment?

Data revealed that 68 per cent of the districts required the evaluator to observe the teacher in her instructional role. There was a wide disparity in the number of observations required, but a sizeable 48 per cent of those using observation techniques as a basis for evaluation required but a single classroom visitation. Interestingly, more districts specified a minimum of three observations than two. One perceives a dual trend for newer policies in this respect. Either they were revised to include additional observations, or the observations were replaced by individual conferences in accompaniment with a self-evaluation model or procedures including performance objectives. All but a few districts who used an observational method stipulated that the evaluator must confer privately with the teacher to discuss the assessment.

Another 10 per cent of the sample reported using individual conferences as the basic procedure in acquiring information relating to a teacher's performance. This procedure was especially evident in districts using a self-evaluation instrument, or in districts where teachers were evaluated on their success in achieving previously established objectives. Conferences apparently are employed only when there is more than a single method utilized in obtaining information concerning a teacher's performance.

Concern seven: In asking for districts' evaluation instruments, this survey sought first, to ascertain the nature of the criteria on which a teacher is assessed, and secondly, to describe generally the various styles and forms of instruments in use.

Teachers, according to the criteria listed in evaluation instruments, in professional growth booklets, or in guidelines distributed to teachers, were assessed most frequently on their instructional skills and personal characteristics. Approximately 75 per cent of the sample evaluated teachers using these criteria. Other criteria that commonly served as a basis for teacher evaluation included, in order of frequency: professional growth, rapport with fellow colleagues and staff, relationships with the community and parents, classroom management ability (if specified separately from instruction), and performance of general school services. Data indicated that pupil achievement was not used by any district as a criterion for assessing teacher performance effectiveness. As might be expected, districts, in revising their evaluation programs to incorporate the assessment of performance objectives, have reduced the prior emphasis placed on a teacher's personal characteristics.

Instruments submitted indicated that teachers continued to be rated in approximately 55 per cent of the sample. Scales either using numbers (1-5) or descriptors (outstanding, excellent, average, weak) were typically employed. There was wide variation as to the number of evaluative criteria listed on the instrument, and the degree to which they were delineated into sub-items. Thus, on numerous instruments, "personal characteristics" might be subdivided to specifically include emotional stability, health, voice, tact, appearance, etc., each of which would be rated. Slightly over one half of the rating instruments provided additional space for the evaluator(s) to elaborate briefly on each item. In approximately 35 per cent of the sample, evaluators, rather than employing a rating instrument, furnished information on a specific comment or general narrative form.

Concern eight: In conclusion, this survey solicited administrators' comments and reactions regarding their districts' current attempts at assessing teacher performance.

Administrators who responded with personal opinions generally expressed reservations about their current practices of teacher evaluation. They conveyed a number of common concerns such as the need, first, for more expertise by the evaluators, and secondly, for sufficient time to conduct a careful and meaningful assessment. Several responses dealt with the intrinsic value of the evaluation process--or lack thereof--as illustrated by this excerpt: ". . . teachers, not wanted in any building, have nothing but positive written evaluations . . ." Many were concerned about increasing the objectivity of the evaluative process by developing procedures for measuring the attainment of performance objectives, thus, placing the greatest emphasis on instruction rather than on the personal attributes of the individual teacher. As one administrator remarked, "The trend appears to be away from rating scales to a process based upon improvement of instruction." Other comments reflected a concern for additional self-evaluation." Other comments reflected a concern for additional self-evaluation procedures and the need, regardless of a district's present program, for continuous reappraisal of the evaluation program itself. In fact, as an optimistic note, a review and/or revision of teacher evaluation programs was reported by 90 per cent of the sample.

CHAPTER IV

SUMMARY AND DISCUSSION

The objectives of this chapter are: (1) to summarize the problem and procedures used in the study; (2) to present conclusions; and (3) to provide a discussion on the relevant concerns regarding teacher evaluation programs.

Summary

The purpose of this study was to gather data and information pertinent to teacher evaluation programs conducted in First Class School Districts in Washington State. Data and information, submitted in response to a mailed questionnaire, dealt with a district's assessment policy, and specifically included concerns relating to evaluative personnel, procedural stipulations, and evaluative instrument or model.

Approximately 60 per cent of the First Class Districts furnished information, which served as the basis for the present study on the status of teacher evaluation programs as currently administered in these districts.

Conclusions

Listed below are the major conclusions of this survey.

1. The majority of first class districts included in the sample reported utilizing a district-wide evaluation model in assessing teacher effectiveness. Data revealed that those districts which encouraged individual school development of evaluative programs tended to be relatively smaller in size.

2. Instructional improvement was cited by most of the sample as the primary purpose for their teacher assessment program. However, several administrators expressed concern over the discrepancy that existed between the procedures and the intended purpose.

3. Principals continued to serve as the primary evaluators of teacher effectiveness. Districts have reported increased use of peers as evaluators, and also in a few isolated instances, the use of pupil-evaluators.

4. Observation was the most frequently used method in obtaining information about a teacher's effectiveness. The survey indicated, however, the newly developed programs, rather than relying solely on observation, included the assessment of performance objectives and self-evaluation techniques such as videotape.

5. The criteria most commonly listed in teacher evaluation instruments is that of personal characteristics and instructional skills. This result is qualified by districts who reported in their revised programs increased emphasis on instructional skills and performance rather than on personal teacher attributes.

6. The rating instrument prevailed as the most common assessment form. In conjunction with this instrument, the majority of districts required a conference with the individual teacher to discuss items deserving commendation or needing improvement.

7. An overwhelming majority of districts reported current teacher evaluation programs under review and/or revision.

Discussion

On the national level, there is concern over the effectiveness of public education. Reflecting this concern, the national assessment, headed by Ralph A. Tyler, is attempting to furnish overall information about the educational attainment of the general population. As the public and educators grow to expect more from education, there is a need to know the status of education, both its strengths and weaknesses. This call for measuring educational effectiveness--referred to as accountability--emphasizes the verifiable measurement of pupil achievement induced as a result of the instructional program. Although its singular intent is to provide information, any assessment of the schools contains broad ramifications.

The incontrovertible trend toward product accountability, which utilizes achievement tests as standard assessment measures, creates many complications for the teacher, currently the primary focus of evaluation procedures. Many teachers feel uneasy about this hard data approach. They point out that achievement tests are not valid indicators of what they have taught, much of which is ephemeral and unretrievable at test time. Teachers, who are unable to control numerous variables including the home environment, a child's readiness, and the specified curriculum, resent being held accountable for learning when they cannot decide what is taught, nor, in many instances, how it is taught. A recent editorial in Saturday Review elaborated on this issue of product accountability:

> As we focus increasingly on pupil performance as a measure of teacher effectiveness, however, it would be easy to forget the complexity of the learning process--that individual children are very different, that they learn different things at different rates, and that even the same child learns at different rates at different times. If, therefore, the laudable effort to improve classroom practice by assessing teacher and school effectiveness merely results in placing more intense and sophisticated pressure on children to perform, the very principle will be denied in practice, for if the concept means anything it is that the ultimate accountability must be to the children (10:41).

Yet a teacher must receive some feedback on whether children learn, otherwise she will not know if she is succeeding, nor in what direction to proceed. The essential questions 'what to measure' and 'how to measure' remain.

Other ramifications of accountability involve the public who assume greater responsibility in specifying explicitly the schools' educational goals. This is no simple task, for parents - and educators--hold widely disparate views on what are appropriate educational goals. Communities also must decide whether the resources provided to the schools are commensurate to the goals they want accomplished.

Thus, many fundamental questions arise as a result of the concern for accountability: "Who is accountable?" "For what are they accountable?" and "Who bears the responsibility for conducting relevant and objective assessment programs?"

These questions are pertinent to our own State of Washington, where the Legislature has reflected the national trend toward accountability by mandating annual evaluation of certificated school personnel. The State law, however, has not resolved the questions surrounding accountability, and teacher evaluation, as indicated by this study, is not equivalent to accountability. While accountability measures product, teacher evaluation has focused on the presage and process variables, such as a teacher's personal characteristics or classroom management skills. While accountability aspires toward objectivity, teacher evaluation has remained essentially subjective, relying upon rating instruments and infrequent observation of the teacher by the evaluator. While accountability provides rigorous and continual reporting of school plans and performances, teacher evaluation has concentrated upon only a singular aspect of education, the role of the teacher.

Too frequently, accountability has focused on the microcosm of the classroom rather than on the macrocosm of the school as a function of the community. In order to hold teachers accountable to the public for their performance, certain prerequisites must be fulfilled. These prior conditions follow a systematic sequence and involve a diversification of the responsibility for the total school program. The prerequisites are:

1. Communities, possibly through group process techniques, need to establish goal statements. Schools have operated too long accepting responsibilities for many of society's frailties. Communities must delineate the appropriate goals or functions for its schools.

2. These goals need to be translated into performance objectives. Teachers will then understand the basis for their assessment, and such objectives will add clarity to their specific assignments.

3. Federal, state and local government agencies must commit themselves to making quality education available to <u>all</u> children, insuring that every individual will learn according to his capacity. This commitment obviously bears a concomitant responsibility to <u>insure</u> that quality education is not denied to pupils because of race, economic station, or financial ability of the local community to support quality education.

4. Methods must be selected or developed by which the performance objectives will be achieved. This requires schools to inform the communities of resources needed to accomplish the established objectives. If resources (teachers, materials, community services, etc.) are not available, then

the objectives should be reconsidered.

5. Measures must be developed in order to accurately appraise attainment of objectives.

6. Procedures need to be established for reporting information to the public in order to plan for subsequent instruction.

This systematic process will require the combined efforts of the schools, communities, and various agencies supporting public education. When each recognizes its specific responsibilities, teacher assessment will assume its proper role as a segment of school accountability intent upon improving instruction.

BIBLIOGRAPHY

1. Anderson, C. P., and S. M. Hunka. "Teacher Evaluation: Some Problems and a Proposal," Harvard Educational Review, XXXIII, 1 (1963), 74-95.

2. Anderson, H. M. "A Study of Certain Criteria of Teaching Effectiveness," Journal of Experimental Education, 23:41-71, September, 1954.

3. Barr, A. S., et al. Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness. 4th Ed. Madison:Dembar Publications, Inc., 1961.

4. Bhaerman, Bob. "A Paradigm for Accountability, "Quest Paper 12. ERIC ED 041 870, August, 1970.

5. Biddle, Bruce T. and William J. Ellena (eds.). Contemporary Research on Teacher Effectiveness. New York: Holt, Rinehart and Winston, 1964.

6. Bowers, Norman D., and Robert S. Soar. "The Influence of Teacher Personality on Classroom Interaction, "Journal of Experimental Education, 30:309-311, June, 1962.

7. Brain, George. "Evaluating Teacher Effectiveness," NEA Journal, 45:35-36, February, 1965.

8. Brighton, Stayner F. Increasing Your Accuracy in Teacher Evaluation. Englewood Cliffs, N. J.:Prentice Hall Inc., 1965.

9. Broudy, Erness B. "A Note on the Validity of the Minnesota Teacher Attitude Inventory (MTAI)," Journal of Educational Research, 64:67.

10. Cass, James. "Accountable to Whom? For What?" Saturday Review, 54:41, March 20, 1971.

11. Clarke, George E. "Teacher Evaluation: Window or Mirror?" The Elementary School Principal in Action. Luther E. Bradfield and Leonard E. Kraft (eds.). Scranton, Pennsylvania: International Textbook Company, 1970.

12. Fattu, Nicholas A. "Effectiveness--an Elusive Quality," Education Digest, 27:24-26, January, 1962.

13. Flanders, Ned A., and Anita Simon. "Teacher Effectiveness," Encyclopedia of Educational Research. Robert Ebel (ed.). New York: Macmillan Co, 1969.

14. Flanders, Ned A. "Categories for Interaction Analysis." n.d. (Mimeo).

15. Gage, N. L. (ed.). Handbook of Research on Teaching. Chicago: Rand McNally & Co., 1963.

16. Gage, N. L. "An Analytical Approach to Research on Instructional Methods," Phi Delta Kappan, 49:601-606, June, 1968.

17. Getzels, J. W., and P. W. Jackson. "The Teacher's Personality and Characteristics," Handbook of Research on Teaching. N. L. Gage ed. Chicago: Rand McNally & Co., 1963.

18. Guba, E. G., and J. W. Getzels. "Personality and Teacher Effectiveness," Journal of Educational Psychology, 46:330-344, October, 1955.

19. Ellena, William J. and Margaret Stevenson, and Harold V. Webb (eds.). Who's a Good Teacher? American Association of School Administrators, Department of Teacher of the NEA, and National School Boards Association, 1961.

20. Harsh, J. Richard. "A Look at Teacher Evaluation." Paper presented at the Annual Conference of the American Association of School Personnel Administrators, Las Vegas, Nevada, October, 1970. ERIC ED 046 993.

21. Heil, Louis M., and Carleton Washburne. "Brooklyn College Research in Teacher Effectiveness," Journal of Educational Research, 55:347-351, May, 1962.

22. Hellfritzch, A. G. " A Factor Analysis of Teacher Abilities," Journal of Experimental Education, 14:166-199, 1945.

23. Hencley, Stephen P. "Impediments to Accountability," Administrator's Notebook, 20 (December, 1971).

24. Kleinman, Gladys S. "Assessing Teacher Effectiveness: The State of Art," Science Education, 50:234-238, April, 1966.

25. Lessinger, Leon. Every Kid a Winner: Accountability in Education. Palo Alto, California: Science Research Associates, 1970.

26. Lins, L. J. "The Prediction of Teaching Efficiency," Journal of Experimental Education, 15:2-60, September, 1946.

27. McPhail, James H. "Teacher Evaluation: A State-by State Analysis." ERIC ED 014 449, October, 1967.

28. MecKlenburger, James A. "Performance Contracting Report," Educational Technology. 11:62, September, 1971.

29. Medley, Donald M., and Harold E. Mitzel. "Measuring Classroom Behavior by Systematic Observation," Handbook of Research on Teaching. N. L. Gage (ed.) Chicago: Rand McNally & Co., 1963.

30. Medley, Donald M., and Harold E. Mitzel. "A Tentative Framework for the Study of Effective Teacher Behavior," Journal of Experimental Education, 30:317-320, June, 1962.

31. Mitzel, Harold E. "Teacher Effectiveness," Encyclopedia of Educational Research. 3rd. ed. Chester W. Harris (ed.). New York: Macmillan, 1966.

32. Mork, G. M. A. "The Verbal Reaction Behavior Log." University of Minnesota, n.d. (Mimeo).

33. Munro, Barry C. "The Minnesota Teacher Attitude Inventory as a Predictor of Teacher Success," Journal of Educational Research, 58:138-139, November, 1964.

34. Ober, Richard L. et al. "An Instructional Manual for the Reciprocal Category System." Based on paper read at the American Educational Research Association, February 8-10, 1968, Chicago, Illinois.

35. Popham, James V., and Eva L. Baker. "A Performance Test of Teaching Effectiveness." ERIC ED 015 144, 1966.

36. Popham, James V. "Objectives '72," Phi Delta Kappan, 53:432-435, March, 1972.

37. Radebaugh, Byron F., and James H. Johnson. "Phase II: Excellent Teachers--What Makes Them Outstanding?" Clearing House, 45:410-418, March, 1971.

38. Ryans, David G. Characteristics of Teachers. Washington, D.C.: American Council on Education, 1960.

39. Ryans, David G. "Teacher Behavior Theory and Research: Implications for Teacher Education, "Journal of Teacher Education, 14:274-93, September, 1963.

40. Scriven, Michael. The Methodology of Evaluation. Lafayette, Indiana: Purdue University, Social Sciences Education Consortiu, 1966.

41. Soar, Rober S. "Methodological Problems in Predicting Teacher Effectiveness," Journal of Experimental Education, 32:287-291, 1964.

42. Stufflebeam, Daniel L. et al. Educational Evaluation and Decision Making. Itasca, Illinois: Fe. E. Peacock Publishers, Inc., 1971.

43. Wick, John W., and Donald L. Beggs. Education for Decision Making in Schools. Boston: Houghton Mifflin Co., 1971.